AD _____

INVESTIGATION OF THE CHARACTER AND PROPERTIES OF ASSUMED SIMILARITY MEASURES

Lee J. Cronbach, Walter Hartmann, Mary E. Ehart
College of Education, University of Illinois

Technical Report No. 7

Study performed under Contract N6ori-07135
with the Office of Naval Research

Project on
Social Perception and Group Effectiveness

February, 1953

INVESTIGATION OF THE CHARACTER AND PROPERTIES OF ASSUMED SIMILARITY MEASURES

Lee J. Cronbach, Walter Hartmann, Mary E. Ehart

In a previous report (2), Fiedler presented evidence that assumed similarity measures are related to team effectiveness in surveying. The present report summarizes the technical analysis of the instrument on which that study was based. Such analysis is a necessary supplement to the validity report, but it has additional scientific value in itself. In particular, we study

(a) the extent to which assumed similarity is a general attitude, and to what extent it depends instead on the content of the test items;

(b) the extent to which assumed similarity measured in four ways (ASp, ASIp, ASo, ASn) reflects the same general quality.

Nature of the Instrument

This project has been concerned with two types of variables which we refer to as Assumed Similarity (AS) and Real Similarity (RS). (Studies of Real Similarity as a predictor of compatibility are to be discussed elsewhere.) The variables are derived from the responses made by subjects under some of the following sets of directions:

s — describe yourself
p — predict how some person you prefer will describe himself
i — describe yourself as you would like to be (ideal)
n — predict how some person you do not prefer (dislike) will describe himself.

The precise directions may be modified in different studies; the ones used in the surveyor study are reported in (2).

By comparing any two sets of responses made by the same person we obtain an AS measure for him. The similarity of his "self" to his perception of his preferred person we denote as ASp, and so on. Figure 1 diagrams the relations yielding AS scores studied in this paper.

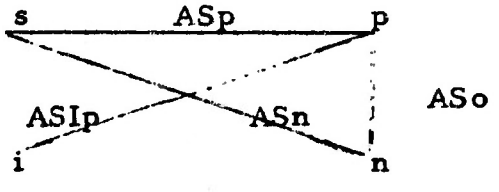

Figure 1. Relations Yielding Certain AS Scores

The resulting split-half coefficients, corrected by the Spearman-Brown formula, are given in Table 1. We discuss them cluster by cluster.

Cluster I. Self-Confidence. The eight items in this cluster included the following:

I am likely to try out my ideas and not worry about the opinion of others.

I often find that I have made up my mind too late. (-)

The direction of scoring was such that the total score represented confidence. The symbol (-) indicates an item reversed in scoring. The cluster appears homogeneous under directions s, p, and i. For some reason, the cluster loses this homogeneity under the n directions. This leads to interesting reflections about the perception of personality. Qualities perceived together, in thinking about the ideal or preferred person, are separated in thinking about the not-preferred person. Such an analysis is irrelevant to our present purpose, and the data at hand are too limited to warrant present pursuit of this lead.

Cluster II. Gregariousness. Naming this trait proved difficult. It was originally intended to be a conventional measure of sociability. The items that survived, however, had a distinctive quality of insulation, or rejection of others. Items include:

I would not want to take another person fully into my confidence. (-)

I am very discriminating in my choice of friends. (-)

This cluster has fairly satisfactory homogeneity, considering its length. The lower reliability for the i directions need not be regarded as significant.

Cluster III. Reflectiveness. This trait is perhaps similar to the Thinking Introversion and Theoretical traits of other tests. Items include:

I try to take an active part in civic and government affairs.

I am more interested in what a person does than in why he does it. (-)

Consistency is relatively low for this scale, probably because items having only peripherally related content were used to lengthen the scale.

Cluster IV. (Doubtful meaning). The low internal consistency for this trait indicates that it was probably assembled on the basis of chance intercorrelations in the Chanute sample on which we had pretested items. It will not be treated further in this report.

Cluster V. Aesthetic Values. Items include:

I can get very affected, perhaps feel like crying, in a sad movie.

Foul language disgusts me.

I am sensitive to color and color schemes.

Table 1

Split-Half Reliabilities, Means and Standard Deviations for Cluster Scores (N = 72)

	Cluster I Self-Confidence			Cluster II Gregariousness			Cluster III Reflectiveness		
	r	Mean	S.D.	r	Mean	S.D.	r	Mean	S.D.
s	.58	36.0	6.9	.51	34.2	6.6	.21	30.5	5.1
p	.46	36.1	5.3	.61	32.6	5.6	.29	31.0	5.8
n	(-.02)*	32.2	4.6	.74	30.2	7.2	.47	30.0	5.7
i	.61	42.6	6.1	.21	35.3	6.1	.50	35.0	5.2

	Cluster IV			Cluster V Aesthetic Values			Median r**
	r	Mean	S.D.	r	Mean	S.D.	
s	(-.07)*	32.0	4.4	.31	32.5	7.0	.31
p	.15	31.6	5.0	.64	31.0	6.7	.46
n	.40	29.7	4.9	.70	31.4	8.6	.47
i	(-.06)*	30.3	4.5	.35	36.4	6.8	.35

*uncorrected

**Cluster IV omitted

Reliabilities for Cluster V ran moderately high. It is strange that in this trait, where people are describing their preferences rather than their actions, the means should be what they are. In this trait the mean ideal score is high, relative to s, p, and n; but the preferred person is said to have less of this trait than either self or the rejected person. The difference between p and i is large enough that it might have arisen on a non-chance basis.

Reading down Table 1, we find that the reliabilities for the four sets of directions are comparable. By this criterion, all directions appear equally usable. It is noteworthy that individual differences in ideal are as distinguishable as differences in self description. It might have been anticipated that, on the contrary, people would so agree in their ideals that no reliable differentiation was possible.

Table 2A presents the intercorrelations of clusters under s directions. The clusters are moderately independent, although I overlaps III, and II overlaps V, in the s and other responses. These intercorrelations will later be compared with intercorrelations of AS measures for the clusters.

Independence of Separate Descriptions and AS Measures

There would be little value in obtaining several descriptions from each man, if the responses of one trial could be predicted from responses under other directions. Table 3 presents the correlations within clusters, and the median over clusters, omitting IV. The correlations are generally positive, except, as expected, that n correlates negatively with the other sorts. The correlation between s and p and that between s and i are so high relative to the reliabilities as to suggest that these three procedures do not provide appreciably different information.

Assumed Similarity in a cluster can be measured by the difference between two cluster scores, ordinarily without regard to sign. Thus ASp_I would be $|s_I - p_I|$. According to the usual theory regarding the reliability of difference scores, these cluster scores are so highly intercorrelated that we could not expect the cluster-scored assumed similarity to be reliable. The AS measures involving the n sort provide an exception. With low correlations between n and s or p, we can expect cluster ASn and cluster ASo to be as reliable as the cluster scores themselves.

When we turn to an examination of AS scored in the usual manner, item by item, we have a quite different picture. To get an item ASp score,

Table 2A

Intercorrelations of Clusters under s Directions

(N = 72)

	I Self-Confidence	II Gregar-iousness	III Reflect-iveness	V Aesthetic Values
I. Self-Confidence	(.58)	.04	.26	.20
II. Gregariousness		(.51)	-.18	-.31
III. Reflectiveness			(.21)	.25
V. Aesthetic Values				(.31)

Table 2B

Median Intercorrelations between Clusters *

(N = 72)

	Median |r|
s directions	.22
p directions	.21
n directions	.09
i directions	.12

* Cluster IV omitted.

Table 3

Intercorrelations of Descriptions within Clusters

(N = 72)

Cluster I
Self-Confidence

	s	p	i	n
s	(.58)	.36	.32	-.21
p		(.46)	.32	-.05
i			(.61)	-.05
n				(.00)

Cluster II
Gregariousness

	s	p	i	n
s	(.51)	.53	-.08	-.27
p		(.61)	.07	-.36
i			(.21)	-.05
n				(.74)

Cluster III
Reflectiveness

	s	p	i	n
s	(.21)	.42	.47	-.20
p		(.29)	.24	-.09
i			(.50)	-.24
n				(.47)

Cluster V
Aesthetic Values

	s	p	i	n
s	(.31)	.48	.64	-.12
p		(.64)	.55	-.01
i			(.35)	-.06
n				(.70)

Median Intercorrelations

	s	p	i	n
s	(.41)	.45	.39	-.20
p		(.53)	.28	-.07
i			(.42)	-.06
n				(.58)

we take the s score on the item, the p score on the item for the same person, and square the difference. These values are summed over items to get an AS score. Taking the square root is required to transform this to the distance scale which is most interpretable, but use of rank correlations makes this transformation unnecessary in our calculations. The split-half reliabilities for AS scores obtained item by item are presented in the diagonals of Table 4, and are extremely high. Obviously, the correlation between descriptions, cluster scored, does not prevent reliable measurement of AS item-by-item. This will be explained below.

The intercorrelations between AS scores are also given in Table 4. ASn and ASo are so highly correlated as to suggest that they will yield little distinct information. One or the other should be dropped from further work, although they might be combined into a single measure. ASp is more independent of the other scores. ASp and ASIp have so much overlap that retention of both in future work seems not likely to be profitable.

Now let us look at the correlation of AS measures from cluster to cluster. ASo was computed for each person on each cluster by two methods. For cluster scoring, the absolute difference, $|p - n|$, on each cluster was used. For item scoring, differences on items were squared and summed, for the items in each cluster. The correlations are given in Tables 5A and 5B. We immediately see that for cluster scoring (Table 5B) the correlations are low. There is a slight trend to positive correlations, suggesting that some overall general quality has very small loadings in all these ASo scores. When we look at ASo item scored, however, we find marked positive intercorrelations even with Cluster IV which is not reliable as a trait measure. For Cluster II a split-half reliability of ASo item scored was found to be .46; the intercorrelations are also of this magnitude.

These high correlations are to be explained, we think, in only one way. This conclusion was originally urged as probable by Fiedler, but seemed improbable to Cronbach and therefore has been tested with considerable care. The data appear to have established beyond question that Assumed Similarity is a mental set or perceptual tendency which influences a person's behavior regardless of the content of the items he is marking. The original aim of cluster scoring was to reduce error. It was argued that error present in p and n ratings would accumulate to a very substantial error in the final ASo score. In cluster scoring, errors in p and n were expected to cancel out to some degree. Yet, it is with "item

Table 4

Rank Order Intercorrelations of AS Measures

(N = 39)

	ASp	ASn	ASo	ASIp
ASp	(.83)	.53	.38	.62
ASn		(.95)	.74	
ASo			(.93)	
ASIp				(.73)

(Note: Diagonal entries are split-half reliabilities (rho).)

Table 5A

Rank Order Correlations between Clusters, Item-Scored

ASo Scores

(N = 35)

	I Self-Confidence	II Gregar-iousness	III Reflect-iveness	IV	V Aesthetic Values
I. Self-Confidence		.48	.59	.41	.32
II. Gregariousness			.40	.44	.18
III. Reflectiveness				.25	.31
IV.					.52
V. Aesthetic Values					

Table 5B

Rank Order Correlations between Clusters, Cluster-Scored

ASo Scores

(N = 35)

	I Self-Confidence	II Gregar-iousness	III Reflect-iveness	IV	V Aesthetic Values
I. Self-Confidence	(.31)	.20	.08	.12	.18
II. Gregariousness		(.56)	.10	.04	-.08
III. Reflectiveness			(.33)	.10	.08
IV.				(.28)	.22
V. Aesthetic Values					(.53)

(Note: Diagonal entries are split-half reliabilities (rho).)

scoring" that appreciable cluster intercorrelations are found. This implies that homogeneity of content within a cluster is of less importance than the reliability of the set to maintain differences between ratings, since a procedure which allows more error in item responses produces a difference score more stable from cluster to cluster.

Scores under one set of directions were correlated with those under another, often nearly to the extent of their reliabilities (Table 3). The reliabilities of AS scores (Table 4) are substantial. An AS score, though, is the result of combining differences between responses under two different directions. The only way one can obtain reliable differences between variables which are themselves correlated to the extent of their reliabilities is to have "correlated error." In other words, subjects tend to have a pervading set to maintain differences between persons as they describe them, regardless of the statements on which this description is made or even regardless of whether the items have any consistency among themselves.

Our analysis does not rule out the possibility that in addition to this general set there may be more specific tendencies. A person might be average in overall ASo, and have greater than average tendency to assume similarity of others in some one trait.

If AS is a set which transcends item content, what does it mean? This we cannot say with any certainty. We have not yet established that a person who shows AS in a particular setting, with a pencil and paper test will show that same set in other relations with persons. The fact that AS scores have been significantly correlated with external criteria in some of Fiedler's studies is evidence that we are dealing with more than a transient verbal set. We are impressed by the possibility that AS represents a tendency to perceive others as alike, or conversely that the person who receives a low AS score tends to be alert to real or fancied differences among others. This would correspond to George Klein's evidence that people differ along a "leveling-sharpening" continuum, as judged by their tendency to differentiate when perceiving laboratory stimuli (6). Further work will be required to test this possible interpretation of AS.

A person who seeks to differentiate may not be an accurate perceiver of others. Gage provides evidence that a description of "people in general" is more likely to fit a particular other person than a description given under directions demanding a deliberate attempt to differentiate and predict him as an individual (5). Hence the person who seeks to differentiate (low AS) may therefore be less accurate.

Conclusions

The instrument in the surveyor study has excellent reliability for measuring Assumed Similarity. ASo determined from differences on the items considered separately has a reliability of .93, ASp of .83, ASn of .95. This method of scoring is superior to a cluster-scoring procedure for determining AS.

The internal consistency of the separate clusters is low. One cluster proved worthless; the other clusters have reliabilities in the neighborhood of .40. The clusters are not intercorrelated to a substantial degree.

Responses obtained under s, p, and i directions are so highly correlated that obtaining separate scores for each of them is not warranted. Similarly, it is found that ASo and ASn are so highly correlated that examining them separately in future studies is not advisable. However, in spite of the high correlation of s with p, it is found that ASp can be measured with good reliability and is partly independent of ASn and ASo.

ASo determined from differences on items within a cluster correlates with ASo in another cluster so highly as to indicate that the AS score does not depend on the content of the items.

Assumed Similarity is therefore to be interpreted as a general attitude, or mental set, essentially independent of the content of the test items. The tendency to assume similarity between persons may be an important aspect of personality. The relation of this trait to other aspects of personality should be studied further.

References

1. Cronbach, L. J., and Gleser, Goldine C. Similarity between persons and related problems of profile analysis. Champaign-Urbana, Illinois: 1952. (Mimeographed, Technical Report No. 2, Contract N6ori-07135 between the University of Illinois and the Office of Naval Research.)

2. Fiedler, F. E. Assumed similarity measures as predictors of team-effectiveness in surveying. Champaign-Urbana, Illinois: 1953. (Mimeographed, Technical Report No. 6, Contract N6ori-07135 between the University of Illinois and the Office of Naval Research.)

3. Fiedler, F. E., Hartmann, W., and Rudin, S. A. The relationship of interpersonal perception to effectiveness in basketball teams. Champaign-Urbana, Illinois: 1952. (Mimeographed, Technical Report No. 3, Contract N6ori-07135 between the University of Illinois and the Office of Naval Research.)

4. Fiedler, F. E., Hartmann, W., and Rudin, S. A. Correction and extension of "The relationship of interpersonal perception to effectiveness in basketball teams." Champaign-Urbana, Illinois: 1953. (Mimeographed, Supplement to Technical Report No. 3, Contract N6ori-07135 between the University of Illinois and the Office of Naval Research.

5. Gage, N. L. Judging interests from expressive behavior. Psychol. Monogr., 1952, 66, No. 18 (Whole No. 350).

6. Klein, G. S. The personal world through perception. In R. R. Blake and G. V. Ramsey (Eds.), Perception - an approach to personality. New York: Ronald Press, 1951. Pp. 328-355.

7. Rudin, S. A., Lazar, I., Ehart, Mary E., and Cronbach, L. J. Some empirical studies of the reliability of social perception scores. Champaign-Urbana, Illinois: 1952. (Mimeographed, Technical Report No. 4, Contract N6ori-07135 between the University of Illinois and the Office of Naval Research.)

8. Warrington, W. G. The efficiency of the Q-sort and other test designs for measuring the similarity between persons. Unpublished doctor's dissertation, Univ. of Illinois, 1952.